

FUR0012-US

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of:

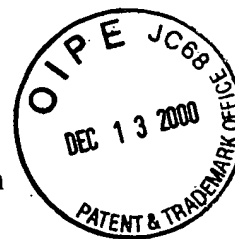
KAZUYUKI ASHIMURA ET AL.

Serial No. 09/671,683

Art Unit: Unknown

Filed: September 28, 2000

Examiner: Unknown



For: SPEECH SYNTHESIS DEVICES

Assistant Commissioner of Patents  
Washington, D.C. 20231

**CLAIM TO PRIORITY UNDER 35 U.S.C. § 119**

Sir:

The benefit of the filing date of the following prior application filed in the following foreign country is hereby requested and the right of the priority provided under 35 U.S.C. § 119 is hereby claimed:

Japanese Patent Appln. No. 11-280528 filed September 30, 1999

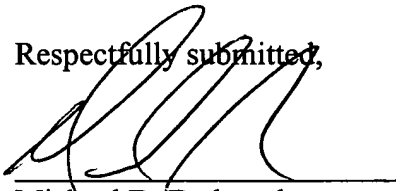
In support of this claim, filed herewith is a certified copy of said foreign applications.

Respectfully submitted,

Date: **December 13, 2000**

SHAW PITTMAN  
2300 N Street, N.W.  
Washington, D.C. 20037-1128  
Tel: (202) 663-8906

By:

  
Michael D. Bednarek  
Reg. No. 32,329

~ 1161568

#4

# 日本国特許庁

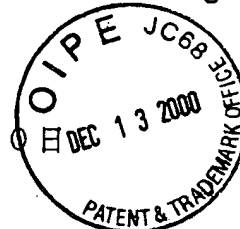
PATENT OFFICE  
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日  
Date of Application:

1999年 9月30日

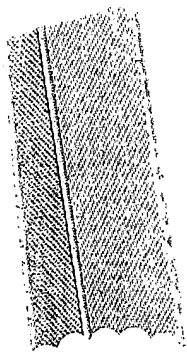


出願番号  
Application Number:

平成11年特許願第280528号

出願人  
Applicant(s):

株式会社アルカディア

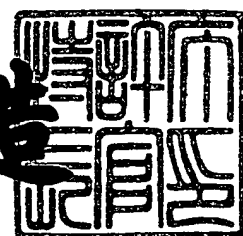


CERTIFIED COPY OF  
PRIORITY DOCUMENT

2000年10月13日

特許庁長官  
Commissioner,  
Patent Office

及川耕造



出証番号 出証特2000-3083595

【書類名】 特許願

【整理番号】 ARK013

【提出日】 平成11年 9月30日

【あて先】 特許庁長官 殿

【国際特許分類】 G01L 3/00

【発明者】

【住所又は居所】 大阪府箕面市坊島 1 丁目 3 番 4 0 号 株式会社アルカデ  
ィア内

【氏名】 芦村 和幸

【発明者】

【住所又は居所】 大阪府箕面市坊島 1 丁目 3 番 4 0 号 株式会社アルカデ  
ィア内

【氏名】 天白 成一

【特許出願人】

【識別番号】 593160323

【氏名又は名称】 株式会社アルカディア

【代理人】

【識別番号】 100092956

【弁理士】

【氏名又は名称】 古谷 栄男

【電話番号】 06-6368-2160

【選任した代理人】

【識別番号】 100101018

【弁理士】

【氏名又は名称】 松下 正

【電話番号】 06-6368-2160

【選任した代理人】

【識別番号】 100101546

【弁理士】

【氏名又は名称】 眞島 宏明

【電話番号】 06-6368-2160

【選任した代理人】

【識別番号】 100106013

【弁理士】

【氏名又は名称】 田川 幸一

【電話番号】 06-6368-2160

【手数料の表示】

【予納台帳番号】 004891

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声合成装置

【特許請求の範囲】

【請求項 1】

人間の発話を収録することによって得たサンプル音声波形データを音声単位に区分するとともに、各音声単位のサンプル音声波形データに対応する音韻情報を関連付けて形成した音声データベースを記録した音声データベース記録手段と、

出力すべき音声の音韻情報を受けて、この音韻情報を音声単位に区分するとともに、前記音声データベースから、音声単位に区分したそれぞれの音韻情報について対応するサンプル音声波形データを取得し、取得した音声単位のサンプル音声波形データを結合して出力すべき音声波形データを得る音声波形合成手段と、

音声波形合成手段によって得られた音声波形データを受けて、アナログ音声信号に変換するアナログ変換手段と、

を備えた音声合成装置であって、

前記音声データベースにおいては、母音を含む音素系列からなっており、複数の音素が明瞭な区分に乏しく連続している場合にはこれら音素を1つのかたまりとして扱った拡張音節に基づいて、サンプル音声波形データを音声単位に区分し

、  
前記音声波形合成手段は、上記拡張音節に基づいて、音韻情報を音声単位に区分することを特徴とする音声合成装置。

【請求項 2】

音韻情報を対応付けたサンプル音声波形データによる音声データベースを用いて、コンピュータに音声合成処理を行わせるための音声合成プログラムを記録した記録媒体であって、

出力すべき音声の音韻情報を受けて、この音韻情報を下記に定義する拡張音節に区分するとともに、前記音声データベースから、拡張音節に区分したそれぞれの音韻情報について対応するサンプル音声波形データを取得し、取得した拡張音節のサンプル音声波形データを結合して出力すべき音声波形データを得る処理をコンピュータに行わせるためのプログラムを記録した記録媒体。

ここで、拡張音節とは、母音を含む音素系列からなっており、複数の音素が明瞭な区分に乏しく連続している場合にはこれら音素を1つのかたまりとして扱ったものをいう。

【請求項3】

出力すべき音声の音韻情報を受けて、この音韻情報を拡張音節に区分する区分手段と、

区分手段によって区分された拡張音節をひとかたまりとして音声波形データを生成し、各拡張音節の音声波形データを結合して出力すべき音声波形データを得る音声波形合成手段と、

音声波形合成手段によって得られた音声波形データを受けて、アナログ音声信号に変換するアナログ変換手段と、

を備えた音声合成装置。

ここで、拡張音節とは、母音を含む音素系列からなっており、複数の音素が明瞭な区分に乏しく連続している場合にはこれら音素を1つのかたまりとして扱ったものをいう。

【請求項4】

コンピュータに音声合成処理を行わせるための音声合成プログラムを記録した記録媒体であって、

出力すべき音声の音韻情報を受けて、この音韻情報を拡張音節に区分し、拡張音節をひとかたまりとして音声波形データを生成し、各拡張音節の音声波形データを結合して出力すべき音声波形データを得る処理をコンピュータに行わせるためのプログラムを記録した記録媒体。

ここで、拡張音節とは、母音を含む音素系列からなっており、複数の音素が明瞭な区分に乏しく連続している場合にはこれら音素を1つのかたまりとして扱ったものをいう。

【請求項5】

音韻情報を受けて、この音韻情報を区分する処理を行うための区分プログラムを記録した記録媒体であって、

音韻情報を受け取り、

当該音韻情報を、下記によって定義される拡張音節に区分する処理をコンピュータに行わせるためのプログラムを記録した記録媒体。

ここで、拡張音節とは、母音を含む音素系列からなっており、複数の音素が明瞭な区分に乏しく連続している場合にはこれら音素を1つのかたまりとして扱ったものをいう。

【請求項6】

サンプル音声波形データを拡張音節に区分して記録した波形データ記録部と、各拡張音節のサンプル音声波形データに対応する音韻情報を関連付けて記録した音韻情報記録部と、

を備えた音声データベースを記録した記録媒体。

ここで、前記拡張音節は、母音を含む音素系列からなっており、複数の音素が明瞭な区分に乏しく連続している場合にはこれら音素を1つのかたまりとして扱ったものをいう。

【請求項7】

音声処理のために用いる音韻情報データを記録した記録媒体であって、

前記音韻情報データは、下記定義による拡張音節をひとかたまりとして扱い、拡張音節ごとに区分情報が付されていることを特徴とする音韻情報データを記録した記録媒体。

ここで、前記拡張音節は、母音を含む音素系列からなっており、複数の音素が明瞭な区分に乏しく連続している場合にはこれら音素を1つのかたまりとして扱ったものをいう。

【請求項8】

音声処理のために用いる音韻辞書を記録した記録媒体であって、

下記定義による拡張音節を単位とする音韻情報に対応づけて、当該音韻の声道伝達関数の時間的变化を記述した音韻辞書を記録した記録媒体。

ここで、前記拡張音節は、母音を含む音素系列からなっており、複数の音素が明瞭な区分に乏しく連続している場合にはこれら音素を1つのかたまりとして扱ったものをいう。

【請求項9】

請求項 1、3 の音声合成装置または請求項 2、4～8 のいずれかの記録媒体において、

前記拡張音節は、母音、母音と長音の結合、母音と二重母音の第 2 要素の結合のいずれかのみを母音要素として含む 1 以上の音素系列であって、最も長いものが優先して拡張音節として選択されるよう定義されることを特徴とするもの。

【請求項 1 0】

請求項 1～9 のいずれかの音声合成装置または記録媒体において、

前記拡張音節は、子音 C（促音、拗音、撥音は含まない）、拗音 y、母音 V（長音、二重母音の第 2 要素は含まない）、長音 R、二重母音の第 2 要素 J、促音 Q、撥音 N を構成要素としたものであって、子音 C、拗音 y の音節量を「0」、母音 V、長音 R、二重母音の第 2 要素 J、促音 Q、撥音 N の音節量を「1」として、各構成要素の音節量合計の多いものが優先して拡張音節として選択されるよう定義されるものをいう。

【請求項 1 1】

請求項 1～10 のいずれかの音声合成装置または記録媒体において、

前記拡張音節には、音節量が「2」の(C)(y)VR、(C)(y)VJ、(C)(y)VNおよび(C)(y)VQを含む重音節と、音節量が「1」の(C)(y)Vを含む軽音節が少なくとも含まれており、

軽音節よりも重音節が優先して拡張音節として選択されることを特徴とするもの。

ここで、(X) は X が含まれていなくとも良く、1 個以上含まれていてもよいことを示す。

【請求項 1 2】

請求項 1 1 の音声合成装置または音声合成プログラムを記録した記録媒体において、

前記拡張音節には、さらに、音節量が「3」の(C)(y)VRN、(C)(y)VRQ、(C)(y)VJN、(C)(y)VJQおよび(C)(y)VNQを含む超重音節が含まれており、

軽音節よりも重音節が、重音節よりも超重音節が優先して拡張音節として選択されることを特徴とするもの。



【請求項 1 3】

請求項 1 の音声合成装置または請求項 2 の音声合成プログラムを記録した記録媒体において、

前記音声データベースは、拡張音節が、その読みを示すかな文字列の長い順に検索可能に構成されていることを特徴とするもの。

【請求項 1 4】

音声波形に対する処理を行う音声処理方法において、

母音を含む音素系列からなっており、複数の音素が明瞭な区分に乏しく連続している場合にはこれら音素を 1 つのかたまりとした拡張音節を分離できない単位として音声波形に対する処理を行う音声処理方法。

【発明の詳細な説明】

【0 0 0 1】

【発明の技術分野】

この発明は、音声合成、音声解析に関するものであり、特にこれら音声処理における処理速度、品質の向上に関するものである。

【0 0 0 2】

【従来の技術および発明が解決しようとする課題】

音声合成の方式としては、規則による合成方式やコーパスベース音声合成方式が知られている。

【0 0 0 3】

規則による合成方式では、与えられた音韻記号列を、音素などの音声単位（「a」「k」など、おおむねローマ字一文字が対応する）に区分し、各音声単位について、基本周波数や声道伝達関数の時間的変化を規則によって決定し、得られた各音声単位の波形を結合して、音声波形として出力するものである。

【0 0 0 4】

しかしながら、各音声単位の波形の結合部分において、不自然さがもたらされることが多かった。また、これを解決するため、音声単位の種類ごとに音声単位と音声単位を結合する際の波形変化等の規則を用意すればよいが、規則の複雑化と処理の低速化を招くこととなり、好ましくなかった。

【0005】

また、コーパスベース音声合成方式では、実際に人間が発話した大量の音声波形とこれに対応する音韻情報を記録した音声のデータベース（音声コーパス）を用意しておき、音声合成の際に、音声コーパスから必要なサンプル音声波形データを切り出して結合することによって、出力すべき音声波形を得るものである。

【0006】

コーパスベース音声合成方式を記述したものとして、匂坂芳典「種類の音韻連接単位を用いた日本語音声合成」電子情報通信学会、1988年3月、ニック・キャンベル他「CHATR：自然音声波形接続型任意音声合成システム」電子情報通信学会、1996年5月、匂坂芳典「コーパスベース音声合成」日本音響学会、1998年11月などがある。

【0007】

これら従来技術におけるコーパスベース音声合成方式では、次のようにして与えられた音記号列に対応する音声波形を得ている。まず、与えられた音韻記号列を音素に区分する。次に、音声コーパス中から、与えられた音韻記号列と最も長く音素列が一致する部分を見いだして、サンプル音声波形を取り出す。取り出したサンプル音声波形を結合して、音声波形を得る。

【0008】

しかしながら、音素を単位として音声コーパスを検索するため、検索処理に膨大な時間を要するという問題があった。また、このように時間を要する割には、最も長く音素列が一致する部分を取り出したにも拘わらず、出力音声が不自然になる場合もあった。

【0009】

そこで、この発明では、上記のような問題点を解決して、処理の迅速性と、出力音声の自然性を両立させて向上させた音声合成装置、音声処理方法を提供することを目的とする。

【0010】

【課題を解決するための手段および発明の効果】

この発明においては、人間の発話が持つ、自然なリズムや、スペクトルのダイ

ナミズムを保存し、より人間らしい音声を合成し、あるいはより精度よく解析を行うために、主として以下の2つの観点から、自然なリズムを保存する音声単位として拡張音節という概念を創設した。

【0011】

観点1：安定した音声波形素片切り出しのための音声単位

観点2：それ以上分割できない、音のリズムの最小単位

拡張音節を音声単位として用いることにより、「母音-母音接続」「母音-半母音接続」や「特殊モーラ」などの、従来、素片接続の連続性に問題のあった箇所において、接続の自然性が改善される。

【0012】

以下、観点1および観点2について説明する。以下合成について説明するが、解析においても同様である。

【0013】

観点1：安定した音声波形素片切り出しのための音声単位

自然な合成音のためには、まず、音声の持つ、スペクトルや基本周波数などの連続量の過渡部において、ダイナミックな動きを音声単位中に保存する必要がある。そのために、音声波形素片の切り出しは、上記の連続量が安定した箇所で行う必要がある。安定した音声波形素片切り出しのための音声単位としては、スペクトルやアクセントの動きを内包したものが好ましい。この出願において発明者が提案する「拡張音節」は、この条件をよく満たしている。

【0014】

観点2：それ以上分割できない、音のリズムの最小単位

話し言葉の自然な合成音声を生成するためには、音声の韻律情報の中でリズムが非常に重要であるため、発話の軸として、まず、リズムを最優先させるべきであると考えられる。

【0015】

話し言葉のリズムは、発話の構成要素である子音や母音の継続時間長の単なる合算によって生じるのではなく、各言語の話者にとってこちよい言語構造が、なんらかの文節単位ごとに繰り返されることによって生じていると考えられる。

例えば、現代日本語の話し言葉では、母音の長さが弁別的であり、長母音や二重母音と単母音では異なる意味を持つため、音声合成にあたって、「長母音（あー）」と「短母音の連鎖（ああ）」の音を互いに流用すると、合成音の品質が損なわれる。

## 【0016】

したがって、発話のリズムをくずさないために、あたかも化学における「分子」のように、「拡張音節」が「リズムの最小単位」として好ましいと考えられる。逆に、発話を「拡張音節」よりも細かく分割してしまうと、音声の持つ自然なリズムがくずれてしまう。

## 【0017】

以上のような観点から、本件出願の発明者は、「拡張音節」という新しい概念を、音声処理に用いたものである。

## 【0018】

この発明の音声合成装置は、人間の発話を収録することによって得たサンプル音声波形データを音声単位に区分するとともに、各音声単位のサンプル音声波形データに対応する音韻情報を関連付けて形成した音声データベースを記録した音声データベース記録手段と、出力すべき音声の音韻情報を受けて、この音韻情報を音声単位に区分するとともに、前記音声データベースから、音声単位に区分したそれぞれの音韻情報について対応するサンプル音声波形データを取得し、取得した音声単位のサンプル音声波形データを結合して出力すべき音声波形データを得る音声波形合成手段と、音声波形合成手段によって得られた音声波形データを受けて、アナログ音声信号に変換するアナログ変換手段とを備え、

前記音声データベースにおいては、少なくとも一つの母音を含む音素系列からなっており、複数の音素が明瞭な区分に乏しく連続している場合にはこれら音素を1つのかたまりとして扱った拡張音節に基づいて、サンプル音声波形データを音声単位に区分し、前記音声波形合成手段は、上記拡張音節に基づいて、音韻情報を音声単位に区分することを特徴としている。

## 【0019】

すなわち、複数の音素が明瞭な区分に乏しく連続している場合にはこれら音素

を1つのかたまりとして扱った拡張音節に基づいて、サンプル波形データから音声単位を取り出すようにしている。したがって、音の特質上区分が困難な部分で、無理矢理にサンプル波形データが結合されるおそれがなく、自然な音声を合成することができる。

## 【0020】

この発明の音声合成装置は、出力すべき音声の音韻情報を受けて、この音韻情報を拡張音節に区分する区分手段と、区分手段によって区分された拡張音節をひとかたまりとして音声波形データを生成し、各拡張音節の音声波形データを結合して出力すべき音声波形データを得る音声波形合成手段と、音声波形合成手段によって得られた音声波形データを受けて、アナログ音声信号に変換するアナログ変換手段とを備えている。ここで、拡張音節とは、母音を含む音素系列からなっており、複数の音素が明瞭な区分に乏しく連続している場合にはこれら音素を1つのかたまりとして扱ったものをいう。

## 【0021】

すなわち、複数の音節が明瞭な区分に乏しく連続している場合にはこれら音素を1つのかたまりとして扱った拡張音節に基づいて、音声合成を行うようにしている。したがって、音の特質上、区分が困難な部分で、無理矢理に合成波形データを結合する必要がなく、自然な音声を合成することができる。

## 【0022】

この発明の音声合成装置は、拡張音節が、母音、母音と長音の結合、母音と二重母音の第2要素の結合のいずれかのみを母音要素として含む1以上の音素であって、最も長いものが優先して拡張音節として選択されるよう定義されることを特徴としている。

## 【0023】

母音と長音の結合、母音と二重母音の第2要素の結合も一つのかたまりとして扱うことにより、自然な音声を合成することができる。

## 【0024】

この発明の音声合成装置は、拡張音節が、子音C（促音、拗音、撥音は含まない）、拗音y、母音V（長音、二重母音の第2要素は含まない）、長音R、二重

母音の第2要素J、促音Q、撥音Nを構成要素としたものであって、子音C、拗音yの音節量を「0」、母音V、長音R、二重母音の第2要素J、促音Q、撥音Nの音節量を「1」として、各構成要素の音節量合計の多いものが優先して拡張音節として選択されるよう定義されることを特徴としている。

【0025】

この発明の音声合成装置は、拡張音節には、音節量が「2」の(C)(y)VR、(C)(y)VJ、(C)(y)VNおよび(C)(y)VQを含む重音節と、音節量が「1」の(C)(y)Vを含む軽音節が少なくとも含まれており、軽音節よりも重音節が優先して拡張音節として選択されることを特徴としている。

【0026】

この発明の音声合成装置は、拡張音節には、さらに、音節量が「3」の(C)(y)VRN、(C)(y)VRQ、(C)(y)VJN、(C)(y)VJQおよび(C)(y)VNQを含む超重音節が含まれており、軽音節よりも重音節が、重音節よりも超重音節が優先して拡張音節として選択されることを特徴としている。

【0027】

この発明の音声合成装置は、音声データベースは、拡張音節が、その読みを示すかな文字列の長い順に検索可能に構成されていることを特徴としている。

【0028】

したがって、音声データベースを順に検索することにより、文字列の長いものを拡張音節として自動的に選択することができる。

【0029】

この発明において、「音声単位」とは、音声合成または解析の際に音声波形をひとかたまりとして扱う単位をいう。

【0030】

「音声データベース」とは、少なくとも音声波形とこれに対応する音韻情報を記録したデータベースをいう。実施形態においては、音声コーパスがこれに該当する。

【0031】

「音声波形合成手段」とは、規則もしくはサンプル波形に基づいて、与えられ

た音韻情報に対応する音声波形を生成する手段をいう。実施形態においては、図 10 のステップ S12～S19、図 17 のステップ S102～S106 がこれに対応する。

#### 【0032】

「プログラム（データ）を記録した記録媒体」とは、プログラム（データ）を記録した ROM、RAM、フレキシブルディスク、CD-ROM、メモリカード、ハードディスク等の記録媒体をいう。また、電話回線、搬送路等の通信媒体も含む概念である。CPU に接続されて、記録されたプログラムが直接実行されるハードディスクのような記録媒体だけでなく、一旦ハードディスク等にインストールした後に実行されるプログラムを記録した CD-ROM 等の記録媒体を含む概念である。さらに、ここでいうプログラム（データ）には、直接実行可能なプログラムだけでなく、ソース形式のプログラム、圧縮処理がされたプログラム（データ）、暗号化されたプログラム（データ）等を含む。

#### 【0033】

##### 【発明の実施の形態】

##### 1. 第 1 の実施形態

##### (1) 全体構成

図 1 に、この発明の一実施形態による音声合成装置の全体構成を示す。この装置は、音声波形合成手段 2、アナログ変換手段 4、音声データベース 6 を備えている。音声波形合成手段 2 は、波形候補取得手段 8、波形候補決定手段 10、波形結合手段 12 を備えている。音声データベース 6 は、人間の発話を収録することによって得たサンプルの音声波形データを、拡張音節に区分して、音韻情報に基づいて検索可能にデータベース化したものである。

#### 【0034】

出力すべき音声の音韻情報は、波形候補取得手段 8 に与えられる。波形候補取得手段 8 は、音韻情報を拡張音節に区分し、音声データベース 6 の中から該当するサンプル音声波形データを取得する。音声データベース 6 には、多くのサンプル波形データが記憶されているので、1 つの拡張音節に対して、複数のサンプル音声波形データが候補として得られる。

【 0 0 3 5 】

波形候補決定手段 1 0 は、波形候補取得手段 8 によって取得された複数のサンプル音声波形データ中から、前後のつながり等を考慮して、1 つの拡張音節に対して 1 つのサンプル音声波形データを決定する。

【 0 0 3 6 】

波形結合手段 1 2 は、波形候補決定手段 1 0 によって得られた一連のサンプル音声波形データを結合し、出力すべき音声波形データを得る。

【 0 0 3 7 】

アナログ変換手段 4 は、これをアナログ音声信号に変換して出力する。このようにして、音韻情報に対応する音声信号を得ることができる。

【 0 0 3 8 】

(2) ハードウェア構成

図 2 に、図 1 の装置を CPU を用いて実現した場合のハードウェア構成の一例を示す。CPU 1 8 には、メモリ 2 0、キーボード／マウス 2 2、フロッピーディスクドライブ (FDD) 2 4、CD-ROM ドライブ 3 6、ハードディスク 2 6、サウンドカード 2 8、A/D 変換器 5 2、ディスプレイ 5 4 が接続されている。ハードディスク 2 6 には、オペレーティングシステム (OS) 4 4 (たとえば、マイクロソフト社の WINDOWS98 など)、音声合成プログラム 4 0 が格納されている。また、音声データベースである音声コーパスを作成するための音声コーパス作成プログラム 4 6 も格納されている。さらに、音声コーパス作成プログラム 4 6 によって作成された音声コーパス 4 2 も格納されている。これらプログラムは、CD-ROM ドライブ 3 6 を介して、CD-ROM 3 8 からインストールされたものである。

この実施形態では、音声合成プログラム 4 0 は、OS と共同してその各機能を実現している。しかし、その一部または全部を、音声合成プログラム 4 0 が単独で実現するようにしてもよい。

【 0 0 3 9 】

(3) 音声コーパスの作成処理

この実施形態による音声合成装置では、音声合成を行う前に、音声コーパス 4



2を作成して用意しておく必要がある。なお、すでに作成された音声コーパス42をハードディスク26にインストールして用いてもよく、また、ネットワーク（LAN、インターネット等）を介して接続された他のコンピュータに格納されている音声コーパス42を用いるようにしてもよい。

#### 【0040】

図3に、音声コーパス作成プログラムをフローチャートにて示す。まず、操作者は、マイク50からサンプルとなる音声を入力する。CPU18は、音声をマイク50から取り込み、A/D変換器52によってデジタルのサンプル音声波形データに変換し、ハードディスク26に記憶する（ステップS1）。次に、操作者は、入力した音声に対応するラベル（音韻情報としての読み）を、キーボード22から入力する。CPU18は、入力されたラベルを、サンプル音声波形データに関連づけてハードディスク26に記録する。

#### 【0041】

図4に、ハードディスク26に記録されたサンプル音声波形データとラベルの例を示す。ここでは、「らいうちゅーいほーが」という音声が入力された場合を例にとって示している。

#### 【0042】

次に、CPU18は、ラベル「らいうちゅーいほーが」を、拡張音節に区分する（ステップS3）。ここで、この実施形態における「拡張音節」とは、母音を含む音のかたまり（音素系列）であって、左最長一致法に基づいて音声単位として切り出したものである。ただし、母音連鎖は、多くとも2つまでを限度とし、母音が3つ連鎖している場合は、2つ目と3つ目の境で区切るようにしている。ここで、「音素」とは、ある一つの言語で用いる音の単位で、意味の相違をもたらす最小の単位である。ある音が当該言語で他の音と弁別的である場合に一つの音素と認められる。

#### 【0043】

図5に、この実施形態による「拡張音節」の構造図を示す。中心となる母音は、単母音（1つの母音）、長母音（母音＋長音）、二重母音（母音＋二重母音の第二要素）のいずれかを必ずとる。その前後に、0個以上の頭子音（子音、拗音

）、尾子音（撥音、促音）が結合したものである。

【0044】

ここで、子音C（促音、拗音、撥音を含まない）、拗音yの音節量を「0」、母音V（長音、二重母音の第二要素を含まない）、長音R、二重母音の第二要素J、撥音N、促音Qの音節量を「1」として、拡張音節の音節量を定義している。すなわち、この音節量にしたがって、重さを規定し、この重さに応じて、拡張音節を3つのタイプに分類している。

【0045】

図6に、この実施形態において用いた「拡張音節」を示す。この実施形態では、「拡張音節」として、音節量「1」の軽音節、音節量「2」の重音節、音節量「3」以上の超重音節を定義している。軽音節は、(C)(y)Vとして示される、「か」「さ」「ちえ」「ぴゃ」などである。いわゆる、モーラと呼ばれるものが該当する。また、(C)は、Cがなくとも、1以上あってもよいことを示す。(y)も同様である。

【0046】

重音節は、(C)(y)VR、(C)(y)VJ、(C)(y)VN、(C)(y)VQとして示される、「とー」「やー」「かい」「のう」「かん」「あん」「ちゅっ」「りゃっ」などである。

【0047】

超重音節は、(C)(y)VRN、(C)(y)VRQ、(C)(y)VJN、(C)(y)VJQ、(C)(y)VNQなどとして示される、「ちえーん」「うーっ」「さいん」「かいっ」「どんっ」などである。

【0048】

図3のステップS3に戻って、CPU18は、拡張音節の定義に従って（定義アルゴリズムまたは拡張音節一覧テーブル等に基づいて）、ラベル「らいうちゅーいほーが」を拡張音節に区分する。なお、この際、CPU18は、もっとも長い拡張音節を、ラベル中から切り出す。したがって、「らい」「う」「ちゅー」「い」「ほー」「が」という7個の拡張音節が切り出される。

【0049】

次に、CPU 18は、図7に示すように、サンプル音声波形70、スペクトログラム（周波数成分の時間的变化）72、拡張音節に区分したラベル74を、ディスプレイ54に表示する。

【0050】

操作者は、この画面を参照しつつ、マウス22を操作して、サンプル音声波形70に区分のマークを付して、拡張音節に区分する（ステップS5）。このようにして、図8に示すように、拡張音節に区分されてラベルの付されたサンプル音声波形（図においては音声ファイル1）がハードディスク26に記録される。

【0051】

次に、CPU 18は、図8に示すようなファイルインデックスを作成し、ハードディスク26に記録する。ファイルインデックスは、拡張音節に区分したラベルと、これに対応するサンプル音声波形データの開始時間と終了時間を記述したものである。なお、各音声ファイルのファイルインデックスの先頭と最後には、始まりと終わりを示すための符号「##」が記述される。ファイルインデックスは、サンプル音声波形データの数だけ生成される。

【0052】

次に、CPU 18は、図9に示す、ユニットインデックスを作成して、ハードディスク26に記録する。ユニットインデックスは、拡張音節のラベルをインデックスにして、サンプル音声波形との対応を付けたものである。たとえば、図9において、見出し「ちゅー」に対応して、拡張音節「ちゅー」のサンプル音声波形が記録されたファイル名「ファイル1」と当該ファイルにおける記録順番「3」が示されている。また、「ファイル2」の「3」番目にも記録されていることが示されている。このように、拡張音節を見出しとして、その拡張音節が記録されている全てのファイルと、当該ファイル内における記録順が記述されたユニットインデックスが作成される。

【0053】

また、ユニットインデックスは、音声合成時の効率的な検索を実現するために、拡張音節のラベルの長さ（かな文字で表記した場合の文字数）により、長い順にソートされて記録される。このようにラベルの長さによってソートすることに

より、結果的に音節量の大きい順にソートされることとなる。

【0054】

以上のようにして、音声コーパス42として音声ファイル、ファイルインデックス、ユニットインデックスが、ハードディスク26に格納される。

【0055】

なお、上記実施形態では、作業者がサンプル音声波形データに対して、区分位置を示すようにしている。しかしながら、波形データの変化や周波数スペクトルの変化等に基づいて、サンプル音声波形データを自動的に拡張音節に区分するようにしてもよい。さらに、CPU18が拡張音節の区分を候補として表示し、作業者が確認または訂正して、サンプル音声波形データを拡張音節に区分するようにしてもよい。

【0056】

#### (4) 音声合成処理

図10、図11に、ハードディスク26に記録された音声合成プログラム40のフローチャートを示す。操作者は、合成音声のターゲット（出力すべき音声）を「かな文字列」としてキーボード22から入力する（ステップS11）。ここでは、ターゲットとして「らいうこーずいけーほーが」と入力されたものとして、説明を進める。

【0057】

なお、この「かな文字列」は、FDD24を介して、フロッピーディスク34から取り込んでもよく、ネットワーク等を介して他のコンピュータから得たものであってもよい。また、かな文字列以外の音韻情報（漢字かな混じり文等）を受けて、ハードディスク26に記録しておいた辞書等によって、「かな文字列」に変換するようにしてもよい。さらに、アクセントやポーズなどの韻律的な情報を付加してもよい。

【0058】

CPU18は、まず、音声コーパス42のユニットインデックスにおける最初の（つまり最長の）見出し（拡張音節）を取得する（ステップS12）。図9によれば、「ちゅー」が取得される。なお、実際のユニットインデックスは、全て

の拡張音節が見出しとされた膨大な量のものであるが、図9では一部のみを示している。

## 【0059】

次に、この拡張音節「ちゅー」が、ターゲットである「らいうこーずいけーほーが」と、左最長一致するかどうかを判断する（ステップS13）。ここでは、一致しないので、ユニットインデックスの次の見出し「こー」を取得し（ステップS14）、同様の判断を行う（ステップS13）。これを繰り返すことにより、拡張音節「らい」において一致することが見いだされる。

## 【0060】

CPU18は、この拡張音節「らい」により、ターゲット「らいうこーずいけーほーが」において、「らい」と「う」の間に拡張音節の区分を行う。つまり、「らい」を拡張音節として切り出す（ステップS15）。このように、文字列の長い順に拡張音節をソートした音声コーパスを用いることにより、効率的に拡張音節の切り出しを行うことができる。

## 【0061】

次に、CPU18は、「らい」のユニットインデックスに基づいて、ファイルインデックスを参照し、図12に示すような候補ファイル（エントリ）を作成する（ステップS15A）。図12では、「らい」の第1候補のファイルが示されている。このファイルには、音声ファイルのファイル名、順番、開始時間、終了時間、ラベルが記録される。候補ファイル（エントリ）は、「らい」についてのサンプル音声波形データの数だけ生成される。

## 【0062】

CPU18は、「らい」について生成したエントリに番号を付して（たとえば、第1候補、第2候補・・・のように）、「らい」に対応づけて記録する（合成ターゲットの音声単位系列の拡張音節候補参照）。図12においては、「らい」について、4つのエントリがあることが示されている。

## 【0063】

上記のようにして、ターゲットから拡張音節を切り出すと、CPU18は、ターゲット中に未処理の部分があるかどうかを判断する。つまり、ターゲット中に

、まだ、拡張音節として切り出されていない部分があるかどうかを判断する（ステップ S 1 6）。

【 0 0 6 4 】

まだ切り出されていない部分があれば、当該部分を対象として、ステップ S 1 2 以下を再び実行する（ステップ S 1 7）。これにより、次に、「う」が切り出されてエントリが生成され、音声単位系列の拡張音節候補が作成される。図 1 2 においては、「う」について、5 つのエントリが生成されている。

【 0 0 6 5 】

以上の処理を繰り返して、拡張音節の切り出しと、これに対応するサンプル音声波形データの特定（つまり取得）が行われる。図 1 2 に、完成した音声単位系列の拡張音節候補を示す。なお、この実施形態では、先頭と最後を示すために、「# #」を記録している。

【 0 0 6 6 】

次に、CPU 1 8 は、複数の拡張音節候補中から、最適候補を決定する（ステップ S 1 8）。この実施形態では、次のような「環境歪み」および「接続歪み」に基づいて、最適候補を決定するようにしている。

【 0 0 6 7 】

ここで、「環境歪み」は、さらに、「ターゲット歪み」と「コンテキスト歪み」を合計したものである。

【 0 0 6 8 】

「ターゲット歪み」とは、ターゲットの拡張音節と、音声コーパスの拡張音節とが一致することを前提とした上で、当該拡張音節の前後の音素環境が一致しない場合に考慮する歪みをいう。ターゲット歪みは、「左方向歪み」と「右方向歪み」を加算したものとして定義する。

【 0 0 6 9 】

「左方向歪み」は、直前の 1 拡張音節が、ターゲットとサンプルで一致する場合「0」とし、一致しない場合には「1」とする。ただし、直前の 1 音素が、ターゲットとサンプルで一致する場合、拡張音節が一致しなくとも「0」とする。さらに、ターゲットの直前の 1 音素が、無音または促音であり、サンプルの直前

の1音素も無音または促音である場合は、一致したものとみなす（つまり「0」とする）。

## 【0070】

「右方向歪み」は、直後の1拡張音節が、ターゲットとサンプルで一致する場合「0」とし、一致しない場合には「1」とする。ただし、直後の1音素が、ターゲットとサンプルで一致する場合、拡張音節が一致しなくとも「0」とする。さらに、ターゲットの直後の1音素が、無音、無声破裂音または無声破擦音であるか、ターゲット自体が促音であり、サンプルの直後の1音素が、無音、無声破裂音または無声破擦音である場合は、一致したものとみなす（つまり「0」とする）。

## 【0071】

「コンテキスト歪み」とは、次の「左方向歪み」と「右方向歪み」を合計したものである。

## 【0072】

「左方向歪み」とは、当該拡張音節を基準として、文頭までの全ての拡張音節が一致している場合、「0」とする。m個目の拡張音節において一致しなくなった場合には、「 $1/m$ 」の歪みとする。

## 【0073】

「右方向歪み」とは、当該拡張音節を基準として、文末までの全ての拡張音節が一致している場合、「0」とする。m個目の拡張音節において一致しなくなった場合には、「 $1/m$ 」の歪みとする。

## 【0074】

「接続歪み」とは、ターゲットにおいて連続する2つの拡張音節（たとえば、「らい」と「う」）に対応する、音声コーパス中の拡張音節候補が、同じ音声ファイルにおいて連続している場合には「0」、そうでない場合には「1」とするものである。すなわち、候補として決定した連続する拡張音節が、音声コーパスにおいても連続している場合には、歪みは生じない。

## 【0075】

CPU18は、上記の「環境歪み」および「接続歪み」の合計が小さく（好ま

しくは最小と) なるように、拡張音節候補を選択する。その選択基準を、図 12 a に模式化して示す。その結果、たとえば、図 13 に示すように、拡張音節候補が選択される。なお、この実施形態では、動的計画法を用いて、好ましい拡張音節候補を決定している。

【0076】

次に、CPU 18 は、上記によって選択された拡張音節候補を結合（接続）して、音声波形データを生成する（ステップ S 19）。この接続の際には、再び、「接続歪み」を考慮する。

【0077】

接続歪みが「0」で連続する複数の拡張音節候補に対しては、そのサンプル音声波形データを、エントリを参照して、音声ファイルから、まとめて取り出す。また、接続歪みが「1」である 2 つの拡張音節候補については、前の拡張音節候補のサンプル音声波形、後ろの拡張音節候補のサンプル音声波形を、それぞれ取り出す。その上で、両サンプル音声波形を接続する。この際、前のサンプル音声波形の終了付近および後ろのサンプル波形の開始付近において、なめらかに接続できる箇所（たとえば、両者の振幅が 0 に近く、かつ、振幅変化の方向が同じ方向であるような箇所）を見だし、当該部分で切り出して接続する。

【0078】

以上のようにして、図 14 に示すような「らいうこーずいけーほーが」に対応する音声波形データが得られる。

【0079】

CPU 18 は、これを、サウンドカード 28 に与える。サウンドカード 28 は、与えられた音声波形データをアナログ音声信号に変換し、スピーカ 29 から音声として出力する。

【0080】

上記実施形態においては、音声コーパスを検索することにより拡張音節を見だして切り出しを行っているが、音声コーパス作成時と同様に、拡張音節の規則に基づいて切り出しを行うようにしてもよい。

【0081】



## (5)その他の実施形態

上記実施形態では、母音連鎖を2以下に限定して拡張音節を定義しているが、母音連鎖が3以上のものを含めてもよい。たとえば、「きゃいーん」「ぎゅおーん」のように、長音および二重母音を含む場合に、これを1つの拡張音節として扱ってもよい。

## 【0082】

なお、母音連鎖を2以下に限定して拡張音節を定義した場合であっても、「接続歪み」が0で連続する複数の拡張音節候補については一つの波形素片としてまとめて切り出すため、一つの波形素片には3以上の母音連鎖が含まれる場合がある。

## 【0083】

また、上記実施形態では、音声コーパスとして、音声波形データを記録している。しかしながら、PARCOR係数などの音響特徴パラメータを記録するようにしてもよい。これにより、音質は劣化するものの、音声コーパスのサイズを小さくすることができる。

## 【0084】

上記実施形態では、図1の各機能をCPUを用いて実現した場合について説明したが、その一部又は全部をハードウェアロジックによって構成してもよい。

## 【0085】

## 2. 第2の実施形態

## (1)全体構成

図15に、この発明の第2の実施形態による音声合成装置の全体構成を示す。この装置は、規則音声合成を行うものであり、区分手段102、音源生成手段104、調音手段106、アナログ変換手段112を備えている。調音手段106は、フィルタ係数制御手段108、音声合成フィルタ手段110を備えている。拡張音節の継続時間長の辞書116には、各拡張音節について、その継続時間長が記録されている。音韻辞書114には、拡張音節ごとに、声道伝達特性の時間的变化が記録されている。

## 【0086】

出力すべき音声の音韻情報は、区分手段102に与えられる。区分手段102は、音韻情報を拡張音節に区分し、フィルタ係数制御手段108、音源生成手段104に与える。また、区分手段102は、拡張音節の継続時間長の辞書116を参照し、区分した各拡張音節の継続時間長を算出する。これを、音源生成手段104に与える。音源生成手段104は、区分手段102からの情報に基づいて、当該拡張音節のための音源波形を生成する。

## 【0087】

一方、フィルタ係数制御手段108は、拡張音節の音韻情報に基づいて、音韻辞書114を参照し、当該拡張音節の声道伝達特性の時間的变化を取得する。フィルタ係数制御手段108は、これに基づいて、当該声道伝達特性を実現するフィルタ係数を音声合成フィルタ110に出力する。したがって、音声合成フィルタ手段110は、与えられた音源波形に対し、各拡張音節に関して時間的同期をとりつつ、声道伝達特性による調音を施し、音声合成波形として出力する。音声合成波形は、アナログ変換手段112によって、アナログ音声信号に変換される。

## 【0088】

## (2)ハードウェア構成

図16に、図15の装置をCPUを用いて実現した場合のハードウェア構成の一例を示す。CPU18には、メモリ20、キーボード/マウス22、フロッピーディスクドライブ(FDD)24、CD-ROMドライブ36、ハードディスク26、サウンドカード28、A/D変換器52、ディスプレイ54が接続されている。ハードディスク26には、オペレーティングシステム(OS)44(たとえば、マイクロソフト社のWINDOWS98など)、音声合成プログラム41が格納されている。これらプログラムは、CD-ROMドライブ36を介して、CD-ROM38からインストールされたものである。また、ハードディスク26には、拡張音節の継続時間長の辞書116、音韻辞書114が記録されている。

## 【0089】

## (3)音声合成処理

図 17 に、音声合成処理プログラムのフローチャートを示す。操作者は、合成音声のターゲット（出力すべき音声）を「かな文字列」としてキーボード 22 から入力する（ステップ S101）。なお、この「かな文字列」は、FDD 24 を介して、フロッピーディスク 34 から取り込んでもよく、ネットワーク等を介して他のコンピュータから得たものであってもよい。また、かな文字列以外の音韻情報（漢字かな混じり文等）を受けて、ハードディスク 26 に記録しておいた辞書等によって、「かな文字列」に変換するようにしてもよい。さらに、アクセントやポーズなどの韻律的な情報を付加してもよい。

#### 【0090】

CPU 18 は、このかな文字列を拡張音節に区分する（ステップ S102）。拡張音節への区分は、拡張音節の定義に基づく規則や拡張音節を列挙したテーブルに基づいて行う。次に、図 18 に示す拡張音節の継続時間長の辞書 116 を参照して、各拡張音節について、その継続時間長を取得する。なお、この辞書を、図 9 のユニットインデックスと同じように、文字の多い順にソートして用意すれば、図 10 のステップ S11～S17 と同様にして、拡張音節の区分と継続時間長の取得を同時に行うこともできる。

#### 【0091】

さらに、CPU 18 は、各拡張音節の文字列、形態素解析によって得たアクセント情報などに基づいて、各拡張音節に対応する音源波形を生成する（ステップ S104）。

#### 【0092】

次に、図 19 に示すような音韻辞書 114 を参照して、各拡張音節に対応する声道伝達関数の時間的な変化を取得する（ステップ S105）。この音韻辞書 114 には、各拡張音節について、声道伝達関数の時間的な変化が記述されている。さらに、各拡張音節の音源波形に対して、上記の声道伝達関数の時間的な変化を実現するように、調音処理（フィルタ処理）を行う（ステップ S106）。

#### 【0093】

このようにして得た音声合成波形を、サウンドカード 28 に与え、音声として出力する（ステップ S107）。

【 0 0 9 4 】

以上のように、この実施形態においては、拡張音節をひとかたまりとして音声合成を行っているので、音声波形の接続部分において不自然さが排除され、品質の高い合成音声を得ることができる。

【 0 0 9 5 】

(4) その他の実施形態

前記第 1 の実施形態において指摘した変更が、この第 2 の実施形態においても同様に適用可能である。

【 0 0 9 6 】

3. その他の実施形態

上記実施形態では、音声合成について拡張音節を用いた場合を説明した。しかしながら、音声処理一般において、拡張音節を基準として処理を行う場合として適用することができる。たとえば、拡張音節をひとかたまりの単位として、音声の解析を行う場合にも適用することができ、解析精度を向上させることができる。

【図面の簡単な説明】

【図 1】

この発明の一実施形態による音声合成装置の全体構成を示す図である。

【図 2】

この発明の一実施形態による音声合成装置のハードウェア構成を示す図である。

【図 3】

音声コーパス作成プログラムのフローチャートである。

【図 4】

サンプル音声波形データとかな文字列を示す図である。

【図 5】

拡張音節の構造を示す図である。

【図 6】

拡張音節の音節量と音節構造との対応関係および拡張音節の例を示す図である。

【図 7】

サンプル音声波形データ、スペクトログラム、拡張音節に区分された文字列を表示した画面を示す図である。

【図 8】

音声ファイルとファイルインデックスとの関係を示す図である。

【図 9】

ユニットインデックスを示す図である。

【図 1 0】

音声合成処理プログラムのフローチャートである。

【図 1 1】

音声合成処理プログラムのフローチャートである。

【図 1 2】

エントリの作成状態を示す図である。

【図 1 2 a】

環境歪みと接続歪みとの関係を示す図である。

【図 1 3】

拡張音節候補の決定を概念的に示す図である。

【図 1 4】

合成された音声波形データを示す図である。

【図 1 5】

第 2 の実施形態による音声合成装置の全体構成を示す図である。

【図 1 6】

第 2 の実施形態による音声合成装置のハードウェア構成を示す図である。

【図 1 7】

第 2 の実施形態による音声合成処理プログラムのフローチャートである。

【図 1 8】

継続時間長の辞書を示す図である。

【図 1 9】

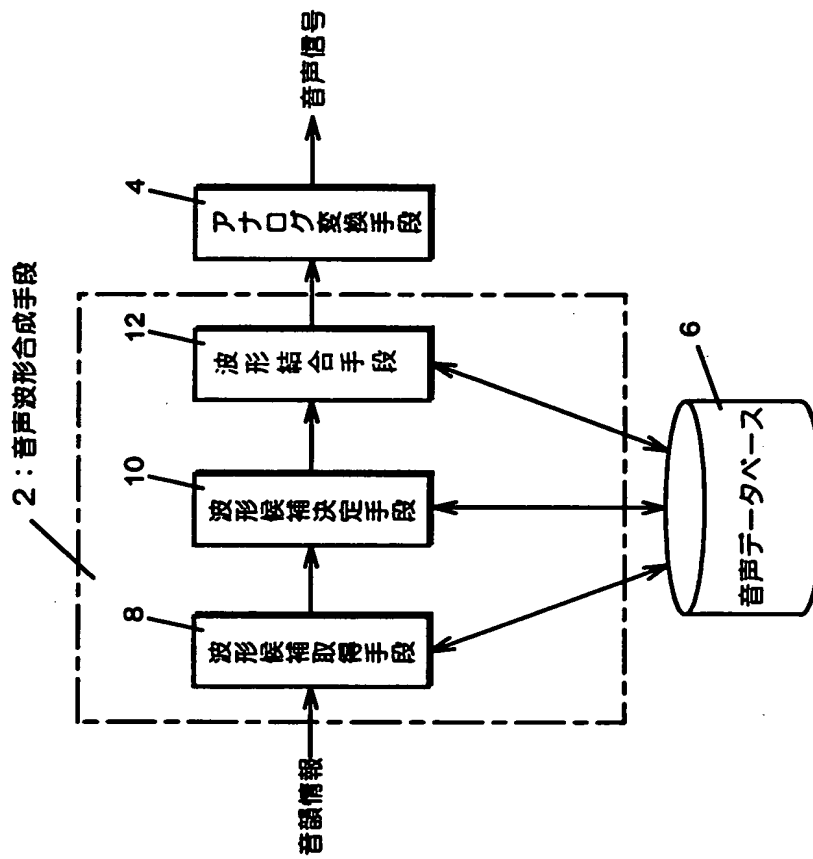
音韻辞書を示す図である。

【符号の説明】

- 4 . . . アナログ変換手段
- 6 . . . 音声データベース
- 8 . . . 波形候補取得手段
- 1 0 . . . 波形候補決定手段
- 1 2 . . . 波形結合手段

【書類名】 図面

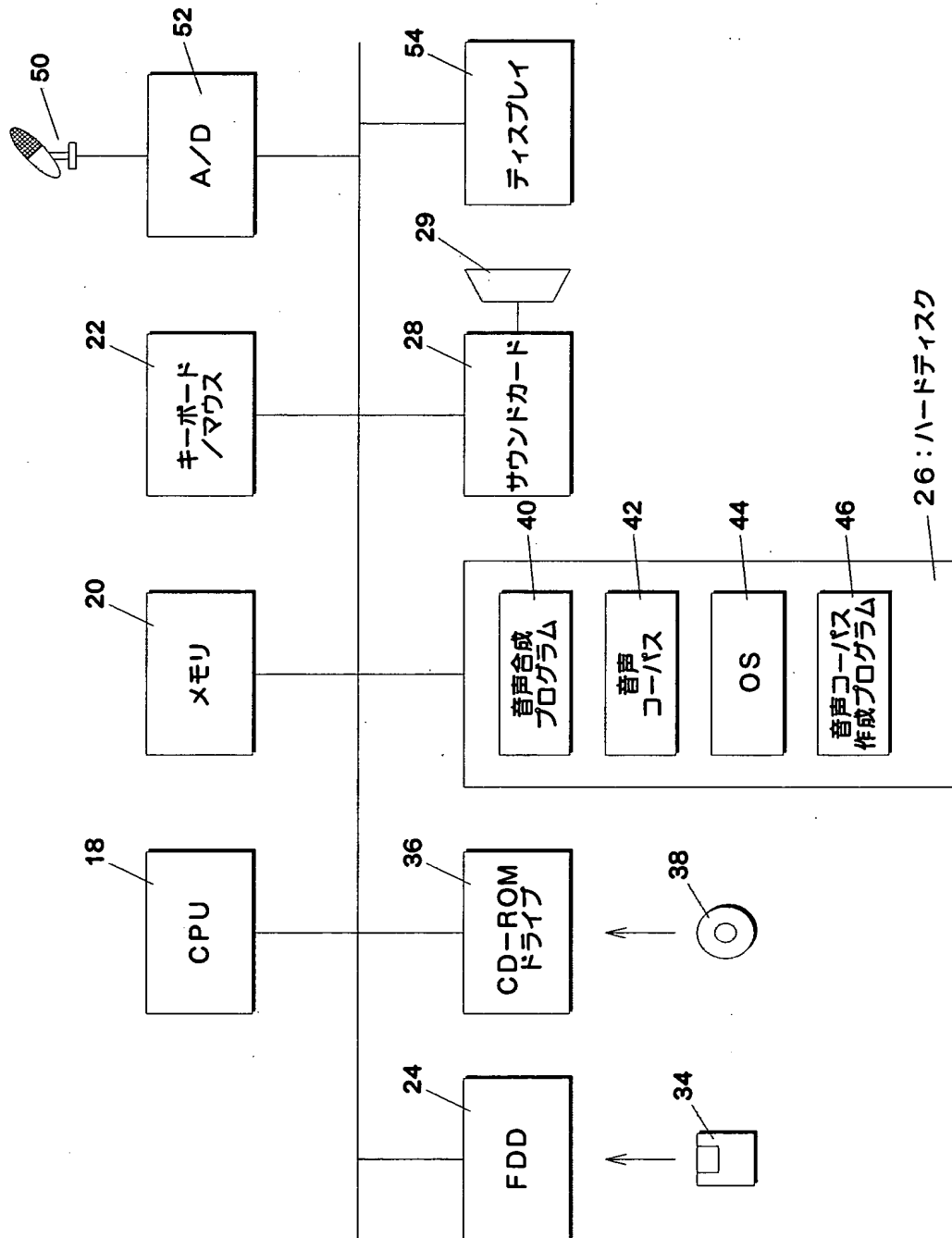
【図 1】



ARK01301

【図 2】

ARK013  
図 2



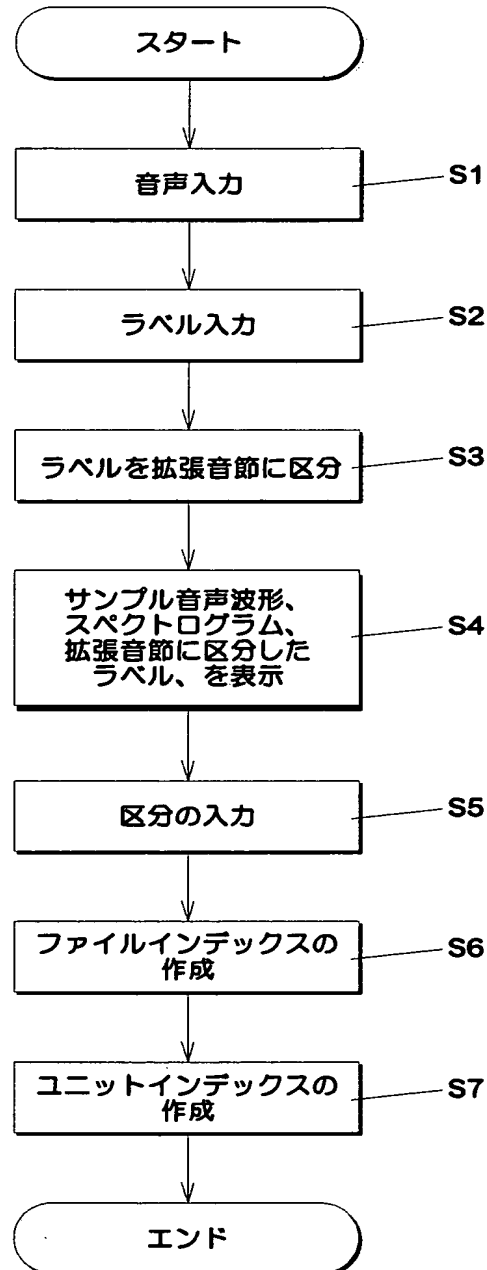
ARK01302



【図 3】

ARK013  
図 3

### 音声コーパス作成プログラム



ARK01303

【図 4】

A



B

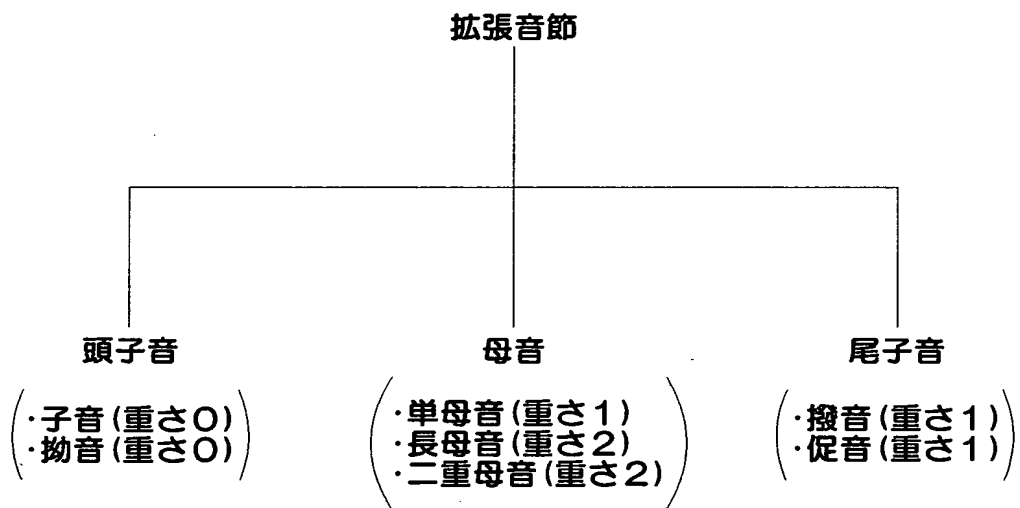
らいうちゅーいほーが

ARK01304

【図 5】

ARK013  
図5

### 拡張音節の構造図



ARK01305

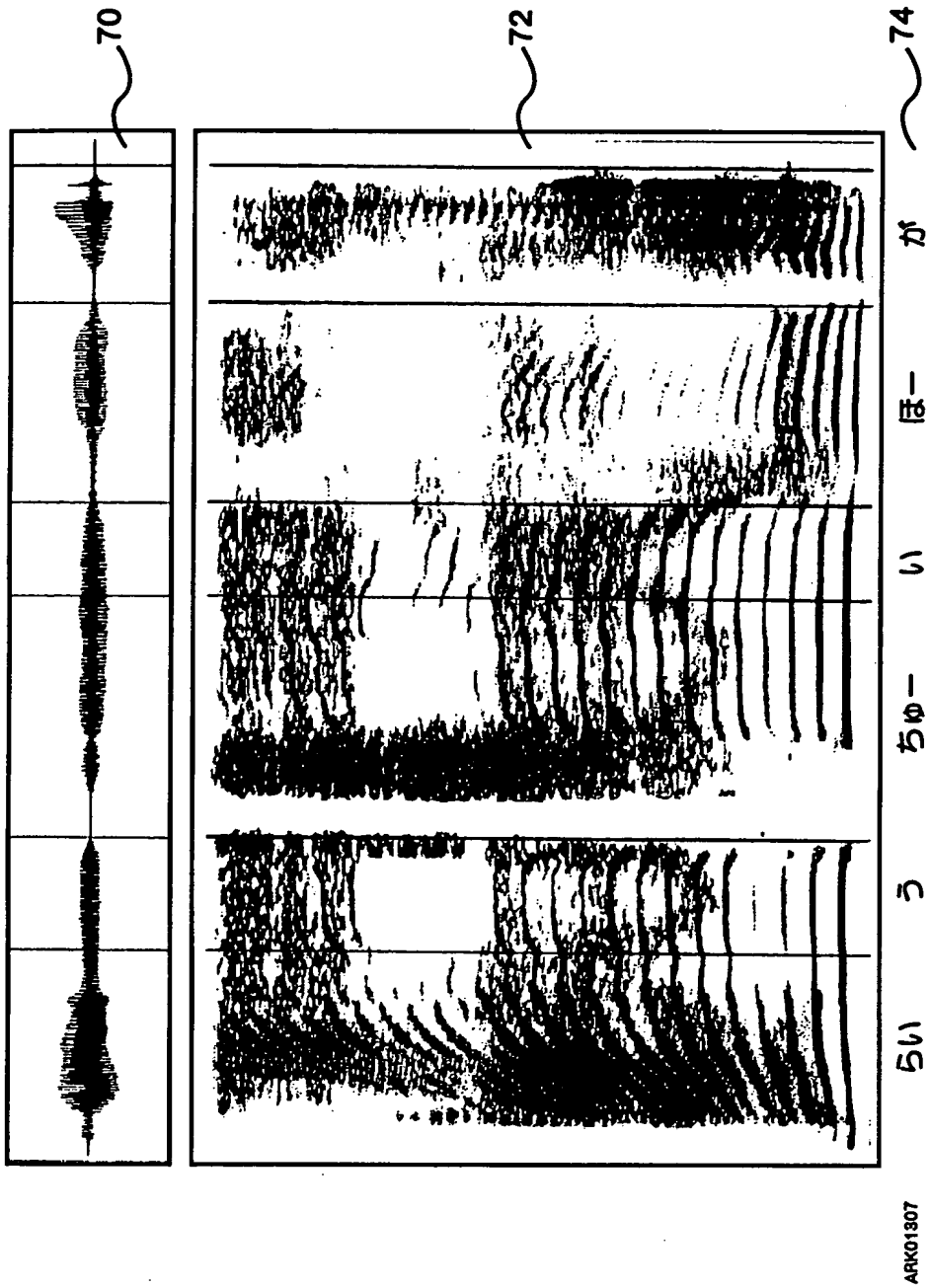
【図 6】

ARK013  
図6

音節量	音節量タイプ	音節構造	例
1	軽音節	(C)(y)V	か, さ, た, な, は, ま, や, ら, … あ, い, う, え, お, ちえ, びゃ, …
2	重音節	(C)(y)VR (C)(y)VJ (C)(y)VN (C)(y)VQ	とー, やー, きゅー, びゅー, … かい, のう, あい, うい, びゅあ, … かん, あん, みゃん, ちゃん, … ちゅっ, りゃっ, じゃっ, ひゅっ, …
3以上	超重音節	(C)(y)VRN (C)(y)VRQ (C)(y)VJN (C)(y)VJQ (C)(y)VNQ など	ちえーん, じゅーん, あーん, … うーっ, ちえーっ, … さいん, ばうん, あうん, ちゃいん, … かいっ, だいっ, きゃいっ, ういっ, … どんっ, うんっ, ちゃんっ, …

ARK01306

【図 7】

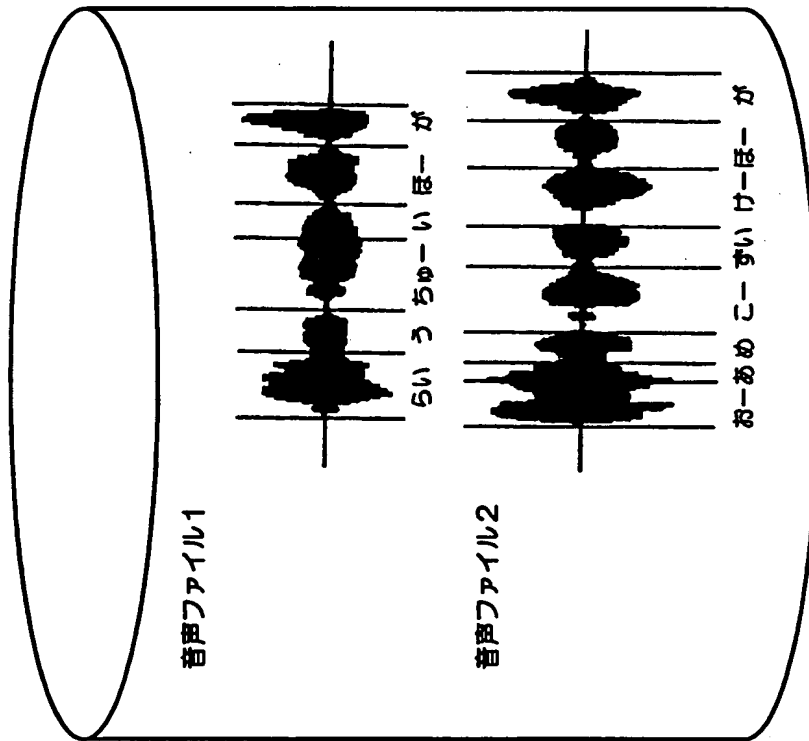


APK01307

【図 8】

ファイルインデックス

音声ファイル3		音声ファイル2		音声ファイル1		開始		終了	
ラベル	開始	終了	開始	終了	開始	終了	開始	終了	
##	0	0	30	605	30				
らい	0	0.330	605	75					
う	0.330	0.505	75						
ちゅー	0.505	0.875							
-	-	-							
-	-	-							
-	-	-							
##	0	0							



ARK01808

【図 9】

ARK013  
図 9

ユニットインデックス

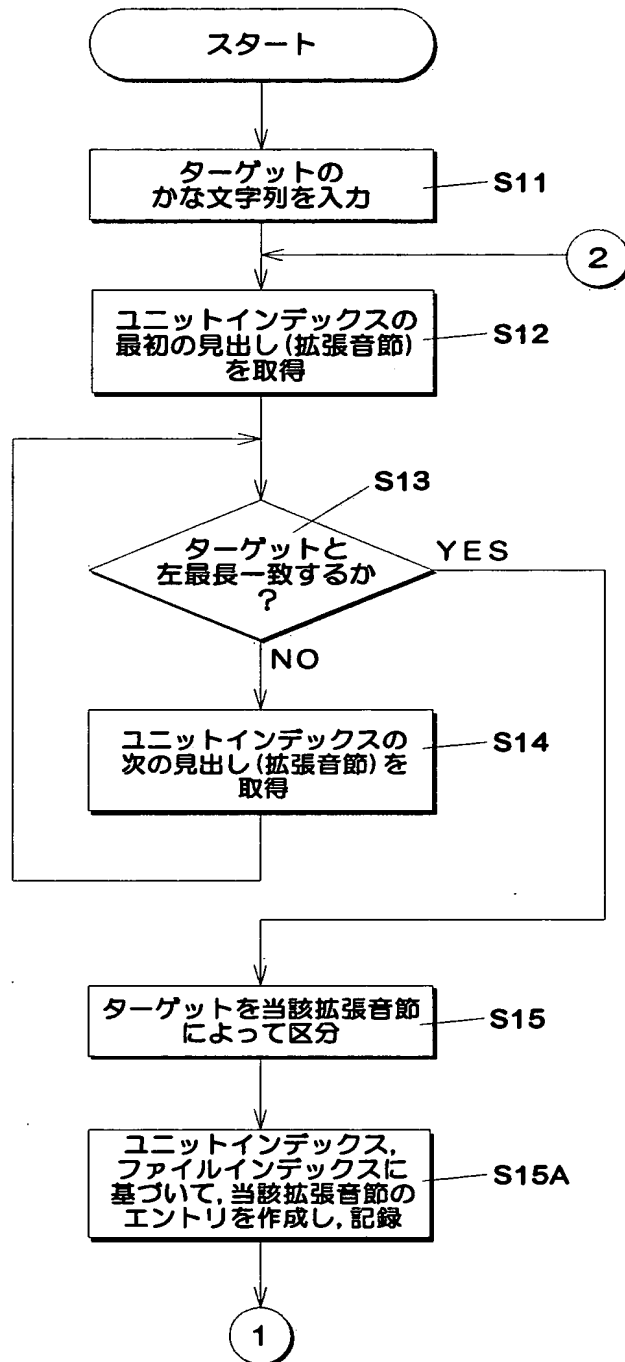
が		ファイル	番号
う		ファイル	番号
らい		ファイル	番号 6
ほー		ファイル	番号 6
ずい		ファイル	番号 6
こー		ファイル	番号 1
ちゅー		ファイル	番号 5
音声ファイル	ファイル	番号 2	5
	ファイル	番号 1	5
	ファイル	番号 3	5
	ファイル	番号 1	3
	ファイル	番号 1	3
ファイル1		番号 3	3
ファイル2		番号 3	3
		番号 2	2
		番号 3	3
---		番号 1	1
		番号 2	2
		番号 3	3
		番号 4	4
		番号 5	5
		番号 6	6

ARK01309

【図 1 0】

ARK013  
図 10

# 音声合成処理プログラム

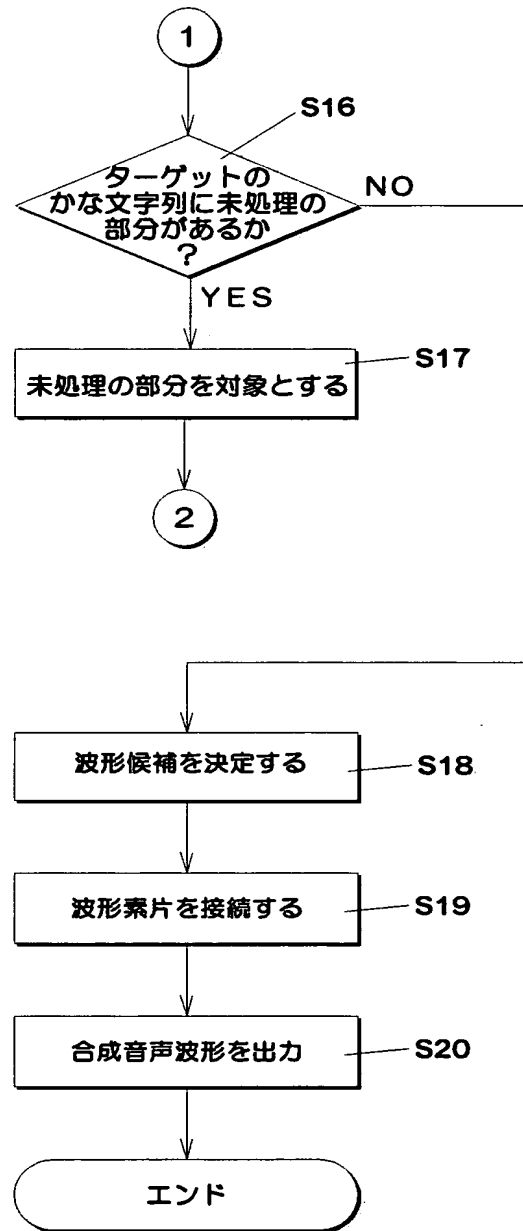


ARK01310



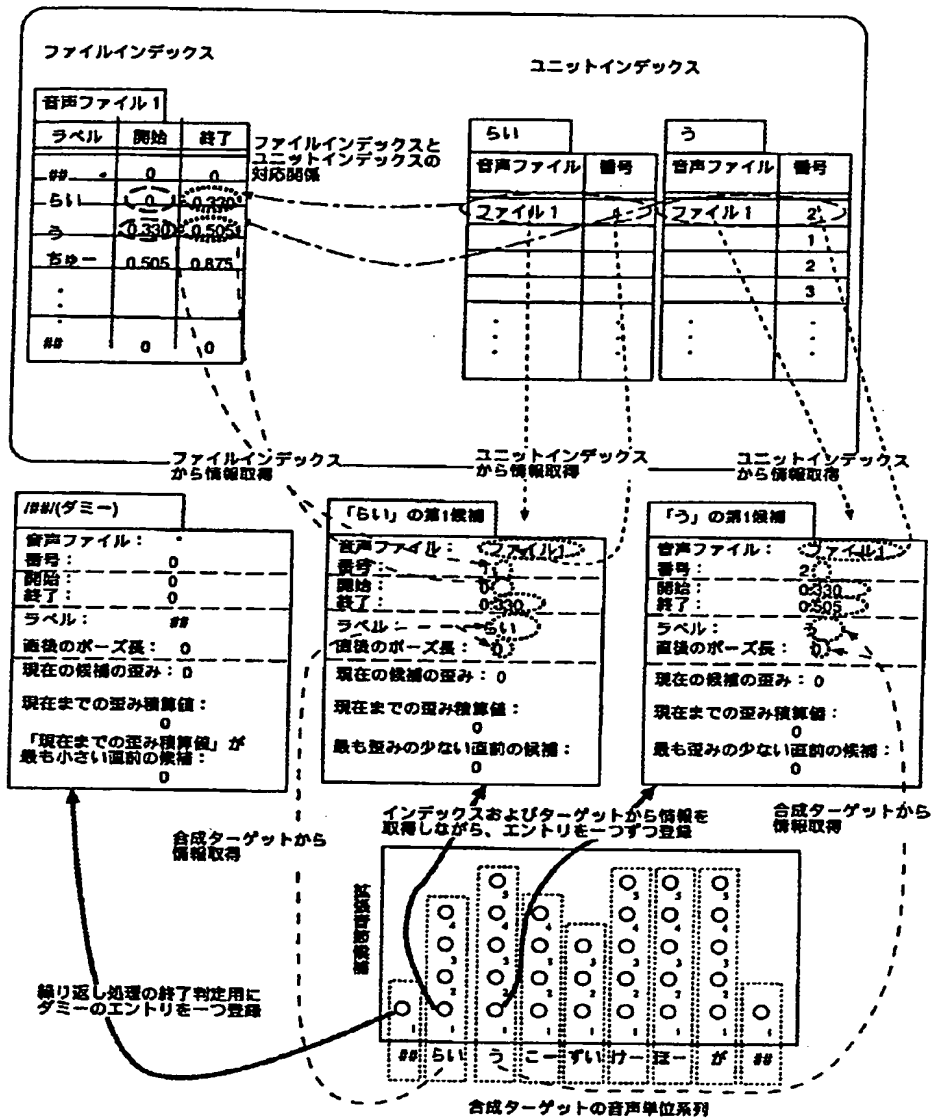
【図 11】

ARK013  
図 11



ARK01311

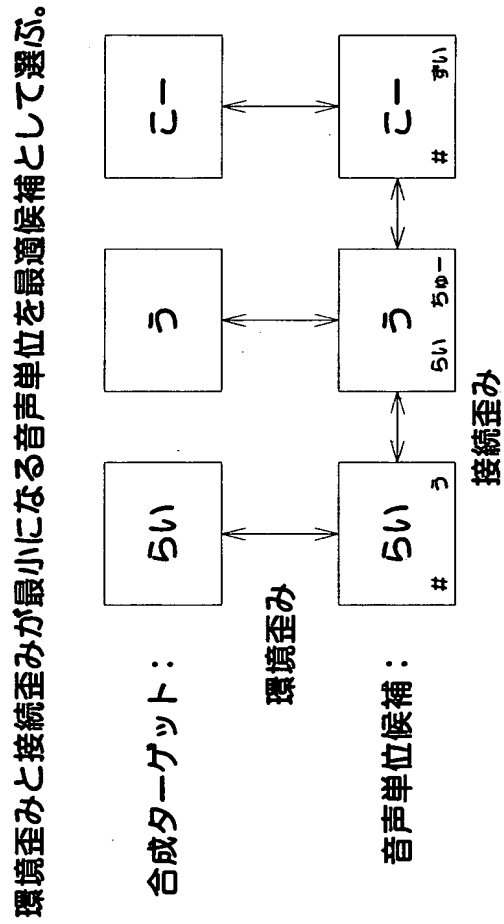
【図 12】



ARK01312

【図 1 2 a】

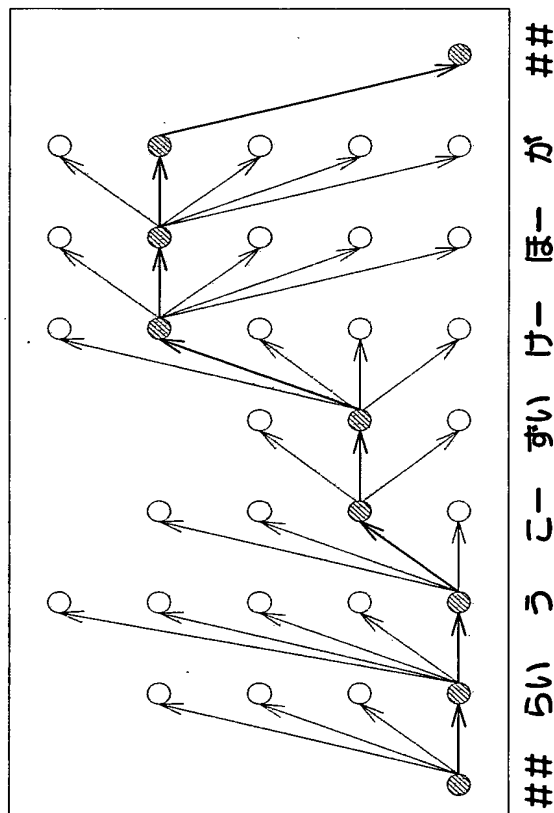
ARKO13  
図 1 2 a



ARKO1312a

【図 1 3】

ARK013  
図 13



拡張音節候補

合成ターゲットの音声単位系列  
(●が、最適な音声単位候補)  
(→が、最適な音声単位候補系列)

ARK01313

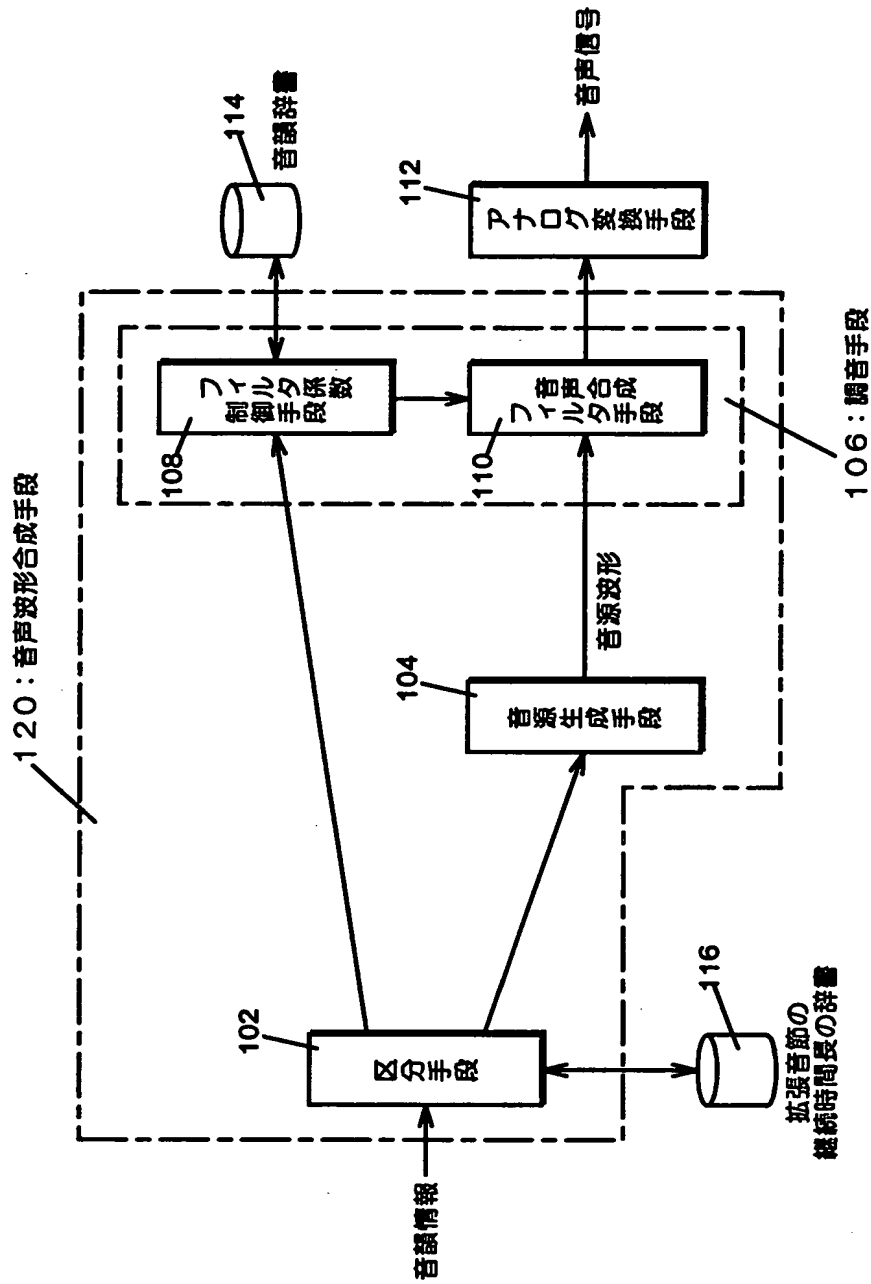
【図 14】



らい う こー すい けー ぽー が

ARK01314

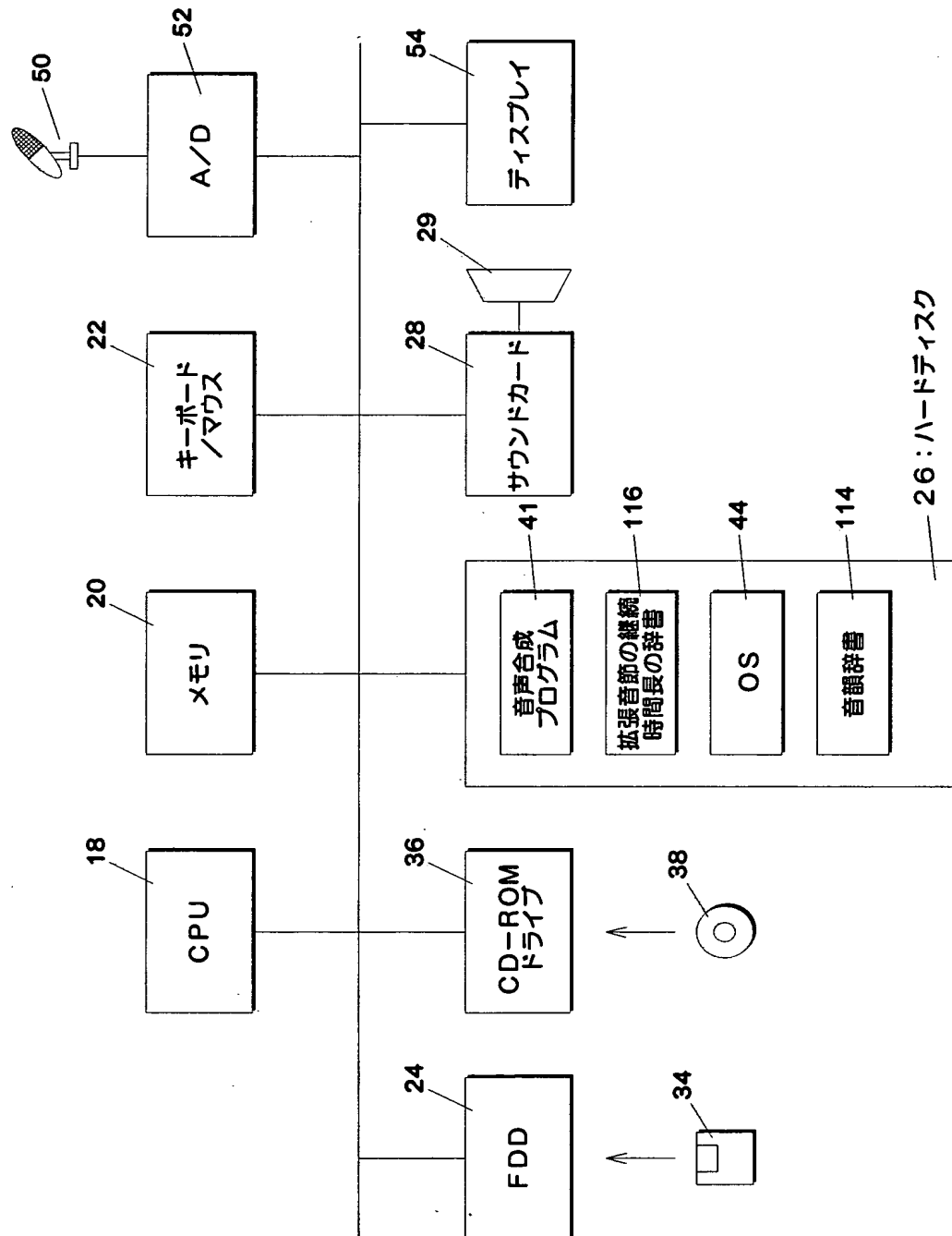
【図 1 5】



ARIK01315

【図 1 6】

ARKO13  
図 16

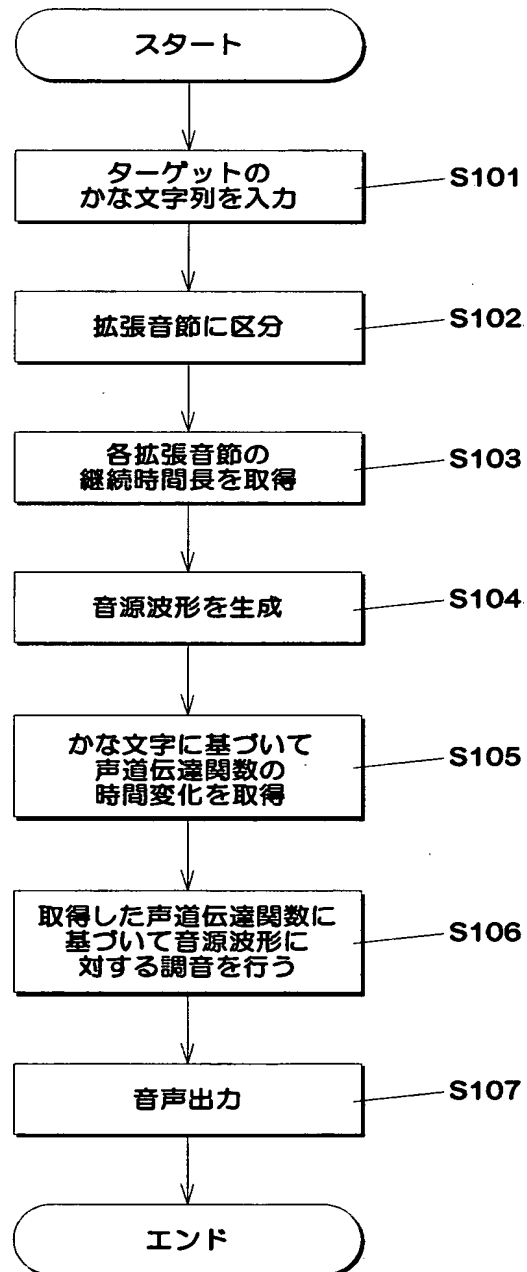


ARKO1316

【図 17】

ARK013  
図 17

### 音声合成処理プログラム



ARK01317



【図 1 8】

ARK013  
図 18

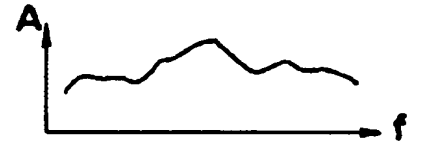
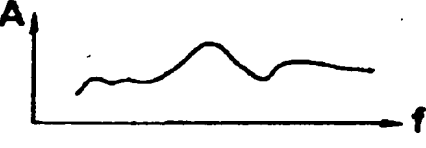

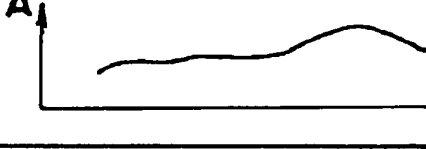
継続時間長の辞書

拡張音節	継続時間長 [ms]
ちゅー	391
こー	213
⋮	⋮
⋮	⋮

ARK01318

【図 19】

音韻辞書

拡張音節	時刻	声道伝達特性
ちゅー	0 } 2 ms	
	⋮	⋮
	383 } 391 ms	
⋮	⋮	⋮
ずい	0 } 3 ms	
	⋮	⋮
	201 } 223 ms	
⋮	⋮	⋮

ARK01319

【書類名】 要約書

【要約】

【課題】 処理の迅速性と、出力音声の自然性を両立させて向上させた音声合成装置を提供する。

【解決手段】 波形候補取得手段 8 は、与えられた音韻情報を拡張音節に区分し、音声データベース 6の中から該当するサンプル音声波形データを取得する。音声データベース 6には、多くのサンプル波形データが記憶されているので、1つの拡張音節に対して、複数のサンプル音声波形データが候補として得られる。波形候補決定手段 10 は、波形候補取得手段 8によって取得された複数のサンプル音声波形データ中から、前後のつながり等を考慮して、1つの拡張音節に対して1つのサンプル音声波形データを決定する。波形結合手段 12 は、波形候補決定手段 10によって得られた一連のサンプル音声波形データを結合し、出力すべき音声波形データを得る。アナログ変換手段 4 は、これをアナログ音声信号に変換して出力する。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [593160323]

1. 変更年月日	1993年 8月30日
[変更理由]	新規登録
住 所	大阪府箕面市坊島1丁目3番40号
氏 名	株式会社アルカディア